

m. Rao



#18

## RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/09/282,879

TIME: 10:50:25

Input Set : N:\Crif3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

## SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Chatterjee, Subroto

7 (ii) TITLE OF INVENTION: RECOMBINANT N-SMASES AND NUCLEIC ACIDS

8 ENCODING SAME

10 (iii) NUMBER OF SEQUENCES: 7

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP

14 (B) STREET: 130 Water Street

15 (C) CITY: Boston

16 (D) STATE: MA

17 (E) COUNTRY: USA

18 (F) ZIP: 02109

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Diskette

22 (B) COMPUTER: IBM Compatible

23 (C) OPERATING SYSTEM: DOS

24 (D) SOFTWARE: FastSEQ Version 1.5

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/282,879

C--> 28 (B) FILING DATE: 31-Mar-1999

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: US/08/774,104

33 (B) FILING DATE:

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Corless, Peter F

37 (B) REGISTRATION NUMBER: 33,860

38 (C) REFERENCE/DOCKET NUMBER: 46906

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: 617-523-3400

42 (B) TELEFAX: 617-523-6440

43 (C) TELEX:

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 1197 base pairs

50 (B) TYPE: nucleic acid

51 (C) STRANDEDNESS: single

52 (D) TOPOLOGY: linear

54 (ii) MOLECULE TYPE: cDNA

55 (iii) HYPOTHETICAL: NO

56 (iv) ANTI-SENSE: NO

W--> 57 (v) FRAGMENT TYPE:

ENTERED

## RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/09/282,879

TIME: 10:50:25

Input Set : N:\Crif3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

```

58      (vi) ORIGINAL SOURCE:
59      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
60      ATGATGACAT ATCACGAAAC GCGCGCGTTG GCTCAAAGCG ACTTACAGCA ACTCTATGCG      60
61      GCACTTGAAA CAACTGAATT TGGCGCTTAC TTTGCGACAC CCGCTGATGA TACTTTACGT      120
62      TTTGGCATTG GCGCAATCGC TACGGCAAAA ACGGCTCAGG CATTACAAGG TGCGGTTGTT      180
63      TTTGGTGC GC AGTCATTTGA TGAACAAGAG TACCCGCAGT CTGAATTGAT GGCGGGTTTT      240
64      TGGTTTGTCC CCGAAGTGAT GGTGACCATC GCGGCAGATA AAATCACGTT CGGATCAGAT      300
65      ACCGTATCTG ATTTTACGAC GTGGCTGGCG CAGTTCGTGC CAAAACAGCC AAATACGGTG      360
66      ACCACTAGTC ATGTGACAGA TGAAGTGGAT TGGATCGAAC GGACAGAGAA TTTGATTGAT      420
67      ACCTTAGCCA TCGATCAAAC CTTAGCCAAA GTCGTTTTTG GTCGGCAACA GACCCTGCAG      480
68      TTATCCGACA CGTTACGACT GGCACAAATT ATTCGTGCGT TAGCTGAGCA GGCGAATACG      540
69      TATCATGTGG TTTTAAAGCG ACATGATGAA TTGTTTATTT CAGCAACACC GGAACGGTTA      600
70      GTGGCTATGT CAGGTGGTCA GATCGCTACG GCGGCGGTCG CTGGGACAAG CCGGCGCGGG      660
71      ACGGATGGCG CTGACGATAT CGCGTTAGGC GAAGCGTTGT TAGCCAGTCA GAAAAACCGC      720
72      ATTGAACATC AATATGTCGT GGCAAGTATC ACGACACGCT TGCAAGACGT GACGACGTCG      780
73      CTAAAGGTGC CGGCCATGCC AAGTTTACTC AAAAATAAGC AAGTTCAGCA TTTGTACACA      840
74      CCAATTACAG GGGACATTGC GGCACATTTA AGTGTGACCG CGATTGTTGA CCGCTTGCAT      900
75      CCAACACCAG CACTGGGTGG CGTCCCACGT GAAGCGGCCC TGTATTACAT TGCGACCCAT      960
76      GAGAAGACAC CTCGTGGCTT GTTTGCAGGT CCTATTGGCT ATTTTACCGC AGATAATAGT      1020
77      GGGGAATTTG TGGTTGGCAT CCGTTCCATG TATGTGAATC AAACGCAGCG ACGAGCAACT      1080
78      TTATTGCTG GTGCCGGGAT TGTGGCTGAC TCCGATGCGC AACAAGAATA TGAAGAACT      1140
79      GGGTTGAAAT TTGAACCCAT GCGGCAATTG TTAAGGACT ACAATCATGT CGAATGA      1197
80
81      (2) INFORMATION FOR SEQ ID NO: 2:
82      (i) SEQUENCE CHARACTERISTICS:
83          (A) LENGTH: 397 amino acids
84          (B) TYPE: amino acid
85          (C) STRANDEDNESS: single
86          (D) TOPOLOGY: linear
87      (ii) MOLECULE TYPE: protein
88      (iii) HYPOTHETICAL: NO
89      (iv) ANTI-SENSE: NO
90      (v) FRAGMENT TYPE: N-terminal
91      (vi) ORIGINAL SOURCE:
92      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
93      Met Met Thr Tyr His Glu Thr Arg Ala Leu Ala Gln Ser Asp Leu Gln
94      1 5 10 15
95      Gln Leu Tyr Ala Leu Glu Thr Thr Glu Phe Gly Ala Tyr Phe Ala
96      20 25 30
97      Thr Pro Ala Asp Asp Thr Leu Arg Phe Gly Ile Gly Ala Ile Ala Thr
98      35 40 45
99      Ala Lys Thr Ala Gln Ala Leu Gln Gly Ala Val Phe Gly Ala Gln Ser
100      50 55 60
101      Phe Asp Glu Gln Glu Tyr Pro Gln Ser Glu Leu Met Ala Gly Phe Trp
102      65 70 75 80
103      Phe Val Pro Glu Val Met Val Thr Ile Ala Ala Asp Lys Ile Thr Phe
104      85 90 95
105      Gly Ser Asp Thr Val Ser Asp Phe Thr Thr Trp Leu Ala Gln Phe Val
106      100 105 110
107      Pro Lys Gln Pro Asn Thr Val Thr Thr Ser His Val Thr Asp Glu Val

```

## RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/09/282,879

TIME: 10:50:25

Input Set : N:\Crif3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

```

115          115          120          125
116 Asp Trp Ile Glu Arg Thr Glu Asn Leu Ile Asp Thr Leu Ala Ile Asp
117          130          135          140
118 Gln Thr Leu Ala Lys Val Val Phe Gly Arg Gln Gln Thr Leu Gln Leu
119          145          150          155          160
120 Ser Asp Thr Leu Arg Leu Ala Gln Ile Ile Arg Ala Leu Ala Glu Gln
121          165          170          175
122 Ala Asn Thr Tyr His Val Val Leu Lys Arg His Asp Glu Leu Phe Ile
123          180          185          190
124 Ser Ala Thr Pro Glu Arg Leu Val Ala Met Ser Gly Gly Gln Ile Ala
125          195          200          205
126 Thr Ala Ala Val Ala Gly Thr Ser Arg Arg Gly Thr Asp Gly Ala Asp
127          210          215          220
128 Asp Ile Ala Leu Gly Glu Ala Leu Leu Ala Ser Gln Lys Asn Arg Ile
129          225          230          235          240
130 Glu His Gln Tyr Val Val Ala Ser Ile Thr Thr Arg Leu Gln Asp Val
131          245          250          255
132 Thr Thr Ser Leu Lys Val Pro Ala Met Pro Ser Leu Leu Lys Asn Lys
133          260          265          270
134 Gln Val Gln His Leu Tyr Thr Pro Ile Thr Gly Asp Ile Ala Ala His
135          275          280          285
136 Leu Ser Val Thr Ala Ile Val Asp Arg Leu His Pro Thr Pro Ala Leu
137          290          295          300
138 Gly Gly Val Pro Arg Glu Ala Ala Leu Tyr Tyr Ile Ala Thr His Glu
139          305          310          315          320
140 Lys Thr Pro Arg Gly Leu Phe Ala Gly Pro Ile Gly Tyr Phe Thr Ala
141          325          330          335
142 Asp Asn Ser Gly Glu Phe Val Val Gly Ile Arg Ser Met Tyr Val Asn
143          340          345          350
144 Gln Thr Gln Arg Arg Ala Thr Leu Phe Ala Gly Ala Gly Ile Val Ala
145          355          360          365
146 Asp Ser Asp Ala Gln Gln Glu Tyr Glu Glu Thr Gly Leu Lys Phe Glu
147          370          375          380
148 Pro Met Arg Gln Leu Leu Lys Asp Tyr Asn His Val Glu
149          385          390          395

```

152 (2) INFORMATION FOR SEQ ID NO: 3:

154 (i) SEQUENCE CHARACTERISTICS:

155 (A) LENGTH: 7 amino acids

156 (B) TYPE: amino acid

157 (C) STRANDEDNESS: single

158 (D) TOPOLOGY: linear

160 (ii) MOLECULE TYPE: protein

161 (iii) HYPOTHETICAL: NO

162 (iv) ANTI-SENSE: NO

163 (v) FRAGMENT TYPE: N-terminal

164 (vi) ORIGINAL SOURCE:

166 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

168 Thr Ser Leu Lys Val Pro Ala

169 1 5

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/282,879

DATE: 09/26/2002

TIME: 10:50:25

Input Set : N:\Crif3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

## 172 (2) INFORMATION FOR SEQ ID NO: 4:

## 174 (i) SEQUENCE CHARACTERISTICS:

175 (A) LENGTH: 7 amino acids

176 (B) TYPE: amino acid

177 (C) STRANDEDNESS: single

178 (D) TOPOLOGY: linear

180 (ii) MOLECULE TYPE: protein

181 (iii) HYPOTHETICAL: NO

182 (iv) ANTI-SENSE: NO

183 (v) FRAGMENT TYPE: N-terminal

184 (vi) ORIGINAL SOURCE:

186 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

188 Arg Ser Ile Thr Val Arg Val

189 1 5

## 192 (2) INFORMATION FOR SEQ ID NO: 5:

## 194 (i) SEQUENCE CHARACTERISTICS:

195 (A) LENGTH: 18 base pairs

196 (B) TYPE: nucleic acid

197 (C) STRANDEDNESS: single

198 (D) TOPOLOGY: linear

200 (ii) MOLECULE TYPE: cDNA

201 (iii) HYPOTHETICAL: NO

202 (iv) ANTI-SENSE: NO

W--&gt; 203 (v) FRAGMENT TYPE:

204 (vi) ORIGINAL SOURCE:

206 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

208 TTGCGGCACT ATTAGGTG

18

## 211 (2) INFORMATION FOR SEQ ID NO: 6:

## 213 (i) SEQUENCE CHARACTERISTICS:

214 (A) LENGTH: 18 base pairs

215 (B) TYPE: nucleic acid

216 (C) STRANDEDNESS: single

217 (D) TOPOLOGY: linear

219 (ii) MOLECULE TYPE: cDNA

220 (iii) HYPOTHETICAL: NO

221 (iv) ANTI-SENSE: NO

W--&gt; 222 (v) FRAGMENT TYPE:

223 (vi) ORIGINAL SOURCE:

225 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

227 CGCCAATGCC AAAACGTA

18

## 230 (2) INFORMATION FOR SEQ ID NO: 7:

## 232 (i) SEQUENCE CHARACTERISTICS:

233 (A) LENGTH: 50 base pairs

234 (B) TYPE: nucleic acid

235 (C) STRANDEDNESS: single

236 (D) TOPOLOGY: linear

238 (ii) MOLECULE TYPE: cDNA

239 (iii) HYPOTHETICAL: NO

240 (iv) ANTI-SENSE: NO

## RAW SEQUENCE LISTING

DATE: 09/26/2002

PATENT APPLICATION: US/09/282,879

TIME: 10:50:25

Input Set : N:\Crf3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

W--> 241 (v) FRAGMENT TYPE:  
242 (vi) ORIGINAL SOURCE:  
244 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
246 GATCCATGAT GACATATCAC GAAACGCGCG TTTCGTGATA TGTCATCATG 50

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/282,879

DATE: 09/26/2002

TIME: 10:50:26

Input Set : N:\Crif3\RULE60\09282879.raw

Output Set: N:\CRF4\09252002\I282879.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:57 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=1  
L:203 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=5  
L:222 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=6  
L:241 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=7